

# SCIENTIFIC REPORTS



OPEN

## Continuous Compressed Sensing for Surface Dynamical Processes with Helium Atom Scattering

Alex Jones<sup>1,\*</sup>, Anton Tamtögl<sup>2,\*</sup>, Irene Calvo-Almazán<sup>2</sup> & Anders Hansen<sup>1</sup>

Received: 02 February 2016

Accepted: 11 May 2016

Published: 15 June 2016

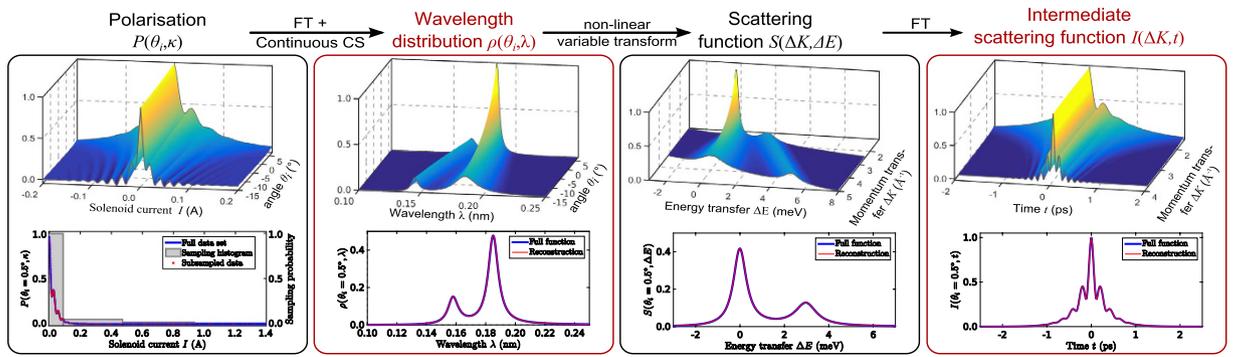
Compressed Sensing (CS) techniques are used to measure and reconstruct surface dynamical processes with a helium spin-echo spectrometer for the first time. Helium atom scattering is a well established method for examining the surface structure and dynamics of materials at atomic sized resolution and the spin-echo technique opens up the possibility of compressing the data acquisition process. CS methods demonstrating the compressibility of spin-echo spectra are presented for several measurements. Recent developments on structured multilevel sampling that are empirically and theoretically shown to substantially improve upon the state of the art CS techniques are implemented. In addition, wavelet based CS approximations, founded on a new continuous CS approach, are used to construct continuous spectra. In order to measure both surface diffusion and surface phonons, which appear usually on different energy scales, standard CS techniques are not sufficient. However, the new continuous CS wavelet approach allows simultaneous analysis of surface phonons and molecular diffusion while reducing acquisition times substantially. The developed methodology is not exclusive to Helium atom scattering and can also be applied to other scattering frameworks such as neutron spin-echo and Raman spectroscopy.

Helium atom scattering has proven to be an invaluable technique to study the structure and dynamics of a wide variety of surfaces ranging from simple metals to reactive and metastable surfaces<sup>1–3</sup>. Helium Spin-Echo (HeSE) spectroscopy is a novel technique<sup>4,5</sup> which combines the surface sensitivity and the inert, completely non-destructive nature of He atom scattering with the unprecedented energy resolution of the spin-echo method<sup>6</sup>. HeSE is the ideal tool for studying surface dynamical processes within a time window from sub-pico second up to nanosecond time scales. This involves atoms and molecules diffusing on the surface<sup>7,8</sup>, phonon vibrations<sup>9,10</sup>, etc. Thanks to the work of Van Hove *et al.*<sup>11</sup> a theoretical framework exists which describes the dynamics of atoms and molecules through the relation of position and time ( $\mathbf{r}, t$ ) to momentum and energy transfer ( $\Delta\mathbf{K}, \Delta E$ ) as a Fourier pair.

In this paper we present a Compressed Sensing (CS) approach for compressing this measurement process, showing that the time needed to reconstruct HeSE spectra can be reduced by several orders of magnitude compared to standard Discrete Fourier Transform (DFT) reconstruction techniques. CS, pioneered by Candès, Donoho, Tao *et al.*<sup>12–15</sup> has long been associated with Nuclear Magnetic Resonance (NMR) based applications such as Magnetic Resonance Imaging<sup>15,16</sup> and NMR Spectroscopy<sup>17</sup>. Recently, compressed sensing has also seen applications focusing on Raman spectroscopy measurements<sup>18</sup> and in molecular dynamics simulations<sup>19</sup>.

Spin-echo spectroscopy shares clear similarities with these fields, such as Fourier transforms arising naturally in data acquisition, however there are also significant differences. In particular, one of the goals of spin-echo spectroscopy is to determine dynamical processes by monitoring the change of polarisation data. Here we consider the whole process of data processing, from polarisation data measurements to the extraction of the molecular dynamics information. Unlike NMR and many other spectroscopy-based applications, after we have performed compression on the initial Fourier transform step we cannot directly use the output data, it must instead undergo several further transforms. Crucially, this includes a non-linear change of variables to momentum/energy space ( $\Delta\mathbf{K}, \Delta E$ ). This precludes the use of standard DFT-based CS techniques as this transform distorts the (necessarily discrete) set of values that we can solve for.

<sup>1</sup>Centre for Mathematical Sciences, University of Cambridge, United Kingdom. <sup>2</sup>Cavendish Laboratory, J. J. Thompson Avenue, Cambridge CB3 0HE, United Kingdom. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.T. (email: tamtoegl@gmail.com)



**Figure 1.** Diagram outlining the various stages of data transformation from measurement in polarisation to the intermediate scattering function (ISF). The upper plots denote the full 2D data while the lower plots are 1D projections/slices with one variable kept constant. The decaying oscillations in the polarisation/ISF are caused by surface phonons. Surface diffusion gives rise to an exponential decay in the polarisation/ISF and a broadening of the peak at  $\Delta E = 0$  of the scattering function. The stages highlighted in red correspond to the target data we wish to reconstruct. The plotted intensities are in consistent arbitrary units (a.u.). Current spacing/resolution is  $2.7 \cdot 10^{-4} \text{A}/10233$  points and 930 points ( $\approx 9\%$ ) are subsampled. The experimental variable  $\kappa$  is proportional to current  $I$  according to (9).

Instead a new *continuous* CS approach, recently introduced by Adcock, Hansen *et al.*<sup>20–22</sup>, is used to reconstruct a continuous approximation that avoids discretising entirely. With this method one has the freedom of evaluating the reconstructed function at any point they desire while still having the speed-up benefits of compressive sampling. Such an approach could be used to handle other inverse problems that require further transformations after reconstruction. Moreover, as demonstrated in refs 20–22 the continuous model avoids the so-called wavelet and inverse “crimes” and hence provides superior reconstructions compared to the classical approach.

Figure 1 shows the process of converting sampled polarisation data to the intermediate scattering function (ISF)  $I(\Delta K, t)$  alongside the application of this new CS technique. Note that CS reconstructions in red for the wavelength intensity and scattering functions match the true signal with 9% of the data traditionally used to reconstruct such spectra using direct Fourier inversion without compression.

Moreover, in standard NMR experiments the smallest group of data that is taken in one measurement is typically a line (or path) of data points in  $k$ -space, with the exception of new realisations such as electron spin echo envelope modulation<sup>23</sup>. However, in HeSE each measurement corresponds to a single point. This gives us an additional degree of freedom in the data acquisition process which makes the application of CS particularly effective, as one can utilize the new approach of structured multilevel sampling in<sup>21</sup> to its fullest to boost performance. In particular, by also taking the structure of the signal into account when designing the sampling strategy one can outperform the classical compressed sensing results (see refs 24 and 25 for experimental validation) that are dictated by the estimate on the number of samples  $m$  to be

$$m \gtrsim s \log(N),$$

where  $s$  is the number of non-zero or important coefficients and  $N$  is the dimension of the vector (The notation  $f(\cdot) \gtrsim g(\cdot)$  means there is a universal constant  $C > 0$  such that  $f(\cdot) \geq Cg(\cdot)$ ).

However, if a signal has  $s = M_1 + s_2 + \dots + s_r$  non-zero coefficients where  $M_1$  denotes the number of the first consecutive non-zero coefficients in the first levels of a wavelet expansion and  $s_j$  is the number of non-zero coefficients in the  $j$ -th level of the wavelet structure then, by using a multilevel sampling procedure<sup>21</sup>, one needs only

$$m \gtrsim M_1 + (s_2 + s_3 + \dots + s_r) \log(N).$$

measurements<sup>21</sup>. Typically, the coefficients corresponding to  $M_1$  are the most important, and most of the energy in the signal is contained in these. This is very convenient as we do not have to pay a log factor for these coefficients. In practice this means substantial gain over the standard approaches as demonstrated in recent publications<sup>24,25</sup>.

### Compressing Spin-Echo Spectra

To understand how continuous CS differs from conventional DFT CS we start with a typical 1D Fourier problem where we measure Fourier samples  $P$  of a wavelength intensity function  $\rho$  we want to reconstruct:

$$P(\kappa) = \int \rho(\lambda) e^{2\pi i \kappa \cdot \lambda} d\lambda, \quad \lambda, \kappa \in \mathbb{R}. \quad (1)$$

From the above it is immediately recognised that  $P$  is a Fourier Transform of  $\rho$  and therefore  $\rho$  can be obtained by the inverse Fourier transform of  $P$ . For a general function  $\rho$  this would require knowing  $P(\kappa)$  at every point  $\kappa \in \mathbb{R}$  which is unrealistic.

In practice however, the wavelength intensity function  $\rho$  is treated as a periodic function

$$\tilde{\rho}(\lambda) = \sum_{l \in \mathbb{Z}} \rho(\lambda + (b-a)l), \quad (2)$$

over a fixed interval  $[a, b]$ . This is convenient because it permits changing the problem to one of handling a Fourier transform to that of handling a Fourier series expansion:

$$\begin{aligned} \tilde{\rho}(\lambda) &= \sum_{l \in \mathbb{Z}} \langle \rho, \chi_{l,\varepsilon} \rangle \chi_{l,\varepsilon}(\lambda), \quad \lambda \in [a, b], \\ \chi_{l,\varepsilon}(\lambda) &= \varepsilon^{1/2} \exp(-2\pi i \varepsilon l \cdot \lambda), \quad l \in \mathbb{Z}, \quad \varepsilon = (b-a)^{-1}, \\ \langle \rho, \chi_{l,\varepsilon} \rangle &= \int_{\mathbb{R}} \rho(\lambda) \overline{\chi_{l,\varepsilon}(\lambda)} d\lambda = P(l\varepsilon)\varepsilon. \end{aligned} \quad (3)$$

The upshot of (3) is that we now only need the values  $P(l\varepsilon)$ ,  $l \in \mathbb{Z}$ , to obtain the wavelength intensity function  $\rho$ , rather than  $P(\kappa)$ ,  $\kappa \in \mathbb{R}$ .

Typically the next step is to truncate the Fourier series expansion, meaning that one makes the approximation

$$\tilde{\rho}(\lambda) \approx \tilde{\rho}_N(\lambda) = \sum_{l=-N}^N \langle \rho, \chi_{l,\varepsilon} \rangle \chi_{l,\varepsilon}(\lambda), \quad \lambda \in [a, b], \quad (4)$$

for some fixed  $N \in \mathbb{N}$ . The problem is now feasible as only finitely many data points  $l$  are required to determine  $\tilde{\rho}_N$ .

Up to this point both continuous CS and conventional DFT CS agree. Conventional DFT CS then breaks up the interval  $[a, b]$  into a uniform grid of  $N$  points

$$\lambda_{j,N} = a + \frac{b-a}{2N+1} j \quad (5)$$

and solves for  $\tilde{\rho}_N(\lambda_{j,N})$ ,  $j = 1, \dots, 2N+1$ . The advantage of doing this is that (4) becomes a vector-matrix equation of the form  $g = Af$  where  $g_j = \tilde{\rho}_N(\lambda_{j,N})$ ,  $A$  is a DFT matrix and  $f$  corresponds to samples of  $P$ . This can be inverted to give  $f = A^{-1}g$  where  $A^{-1}$  is still a DFT matrix, and therefore an isometry, which facilitates the application of CS.

The drawback of this approach is that important information could be lost by only considering  $\tilde{\rho}_N(\lambda_{j,N})$ ,  $j = 1, \dots, 2N+1$  instead of (4). The Fourier series approximation could have been evaluated at any point in the interval  $[a, b]$  and suddenly the best we can do is reconstruct the  $\tilde{\rho}_N(\lambda_{j,N})$ , even though we are still working with the same number of Fourier samples. Do we really have to pay this price in order to be able to compress this problem? The answer is: no.

Motivated by the Fourier series (4), one can try approximating  $\rho$  in terms of a new *Reconstruction Basis*  $\sigma_n$ ,  $n \in \mathbb{N}$

$$\tilde{\rho}(\lambda) \approx \sum_{n=1}^M \langle \rho, \sigma_n \rangle \sigma_n(\lambda), \quad \lambda \in [a, b]. \quad (6)$$

Apart from the benefit of keeping the problem continuous, one also has the freedom to *choose* which basis  $\sigma_n$  to work with, making the approach more versatile than a straight DFT approach. Note that this is similar in spirit to the work on finite rate of innovation<sup>26</sup> as well as to the concept suggested by Markovich *et al.*<sup>27</sup>.

Since one is still sampling data that corresponds to Fourier coefficients of  $\rho$ , it is impossible to exclusively work with their choice of basis  $\sigma_n$ . Instead one has to convert Fourier series coefficients into coefficients in the basis  $\sigma_n$ . This is achieved by working with the infinite change of basis matrix for the two bases:

$$B_{l,n} = \langle \sigma_n, \chi_{l,\varepsilon} \rangle, \quad n \in \mathbb{N}, \quad l \in \mathbb{Z}. \quad (7)$$

Using this matrix to reconstruct the coefficients  $\langle \rho, \sigma_n \rangle$ ,  $n = 1, \dots, M$  we can then use (6) to approximate  $\tilde{\rho}(\lambda)$ . As long as we assume the  $\sigma_n$  form an orthonormal basis then the matrix  $B$  is an infinite-dimensional isometry which also allows the application of CS thanks to the work of refs 20 and 22.

This continuous approach has two significant advantages:

- The approximation  $\tilde{\rho}$  is now a continuous function (as opposed to discrete) that can be evaluated at any point and hence this allows the non-linear change of variables going from the wavelength distribution to the scattering function  $S$  (as shown in Fig. 1). Such a transform is not possible with conventional discrete CS techniques.
- The approximation  $\tilde{\rho}$  is computed with the actual coefficients in the new expansion of the wavelength function, and hence this approximation has the characteristics of the approximation in the new basis rather than the truncated Fourier series. This means reducing Gibbs ringing and other artefacts coming from Fourier approximations (see refs 20–22 for details).

Details on the theoretical background of CS, what basis to use, the convex optimisation problems we solve and how to subsample the Fourier data are provided later on in the paper.

**Paper Outline.** After a short introduction into HeSE spectroscopy and compressed sensing/continuous compressed sensing we focus on the first Fourier transform step shown in Fig. 1. CS is used to demonstrate the compressibility of phonon detection.

In Surface Diffusion Using Continuous CS, we consider the full cycle of transforms shown in Fig. 1, from polarisation data to the determination of diffusion processes. Thereafter we elaborate on the advantages of the continuous CS technique in order to measure processes that appear on different energy scales.

## Helium Spin-Echo Spectroscopy

The principle of the Helium spin-echo apparatus is the following: A beam of thermal  $^3\text{He}$  is generated from the source in a fixed direction. The nuclear spins are polarised and then rotated by the initial or incoming solenoid before being scattered from the target crystal surface. Afterwards any scattered He atoms heading in the direction of the detector are then rotated by the final or outgoing solenoid and passed through another polarisation filter. Thereby the apparatus achieves an energy resolution of  $3\mu\text{eV}$  and dynamical processes within a time window spanning from the sub-picosecond regime up to nanoseconds can be observed. While a schematic sketch of the machine used at the Cavendish laboratory can be found in the Supplementary Information further details can be found in Jardine *et al.*<sup>4</sup>.

Key variables that the operator can freely adjust include:

- The currents  $I_i, I_f$  that run through the initial and final solenoid respectively.
- The scattering geometry, namely the angle of the surface normal relative to the source/detector setup.

The incoming monochromatic He beam can be viewed as a plane wave with propagation wavevector  $\mathbf{k} \in \mathbb{R}^3$  and angular frequency  $\omega$ :

$$\psi(\mathbf{r}, t) = \exp(i(\mathbf{k} \cdot \mathbf{r} - \omega t)), \quad \mathbf{r} \in \mathbb{R}^3, \quad t \in \mathbb{R}. \quad (8)$$

Here,  $\mathbf{r}$  denotes position and  $t$  is time. The wavevector  $\mathbf{k}$  and wavelength  $\lambda$  are related to particle's momentum  $\mathbf{p}$  by the de Broglie relations via  $\mathbf{p} = \hbar\mathbf{k}$  and  $|\mathbf{p}| = p = 2\pi/\lambda$ . Furthermore, the frequency  $\omega$  is related to the particle energy  $E$  by the relation  $E = \hbar\omega$ . Using these relations we can treat  $\mathbf{k}$  as representing momentum and  $\omega$  as energy. Notice that by Formula (8) we have identified two Fourier pairs  $(\mathbf{k}, \mathbf{r}), (\omega, t)$ .

Upon scattering from a dynamic surface, the wavevector  $\mathbf{k}$  and the energy of the He atom  $E$  before and after the scattering will typically change. By measuring the probability of a He atom to go from an initial state  $i$  to a final state  $f$  information about the surface dynamical processes can be gained. Therefore, we need to measure the properties of the initial and final He beam which is in practice done by changing the current through the solenoids.

In the following we show how the solenoid currents are related to the scattering wavelengths  $(\lambda_i, \lambda_f)$ . This section follows closely the review of Alexandrowicz and Jardine<sup>5</sup> and a more detailed description can be found in the Supplementary Information.

**Solenoid Currents and Spin Polarisation.** The solenoid currents  $(I_i, I_f)$  and scattering wavelengths  $(\lambda_i, \lambda_f)$  share a direct Fourier relationship. Recall that we have two solenoids that generate magnetic fields which rotate the polarisation of the He beam. The solenoid current determines the strength of the magnetic field but it is more convenient to use the experimentally controllable parameter  $\kappa$  which is proportional to the current in the solenoids via:

$$\kappa_i = \frac{m\gamma B_{\text{eff}} I_i}{2\pi\hbar}, \quad \kappa_f = \frac{m\gamma B_{\text{eff}} I_f}{2\pi\hbar}, \quad (9)$$

where  $\gamma$  is the gyromagnetic ratio of the He atom,  $m$  is its mass and  $B_{\text{eff}}$  is an apparatus specific constant. The polarisation of the He beam in terms of amplitude and phase can be conveniently written as a complex number. When using the scaled variables of (9) the measured polarisation of the He beam in the detector can be represented as the two-dimensional Fourier transform of  $\rho(\lambda_i, \lambda_f)$ :

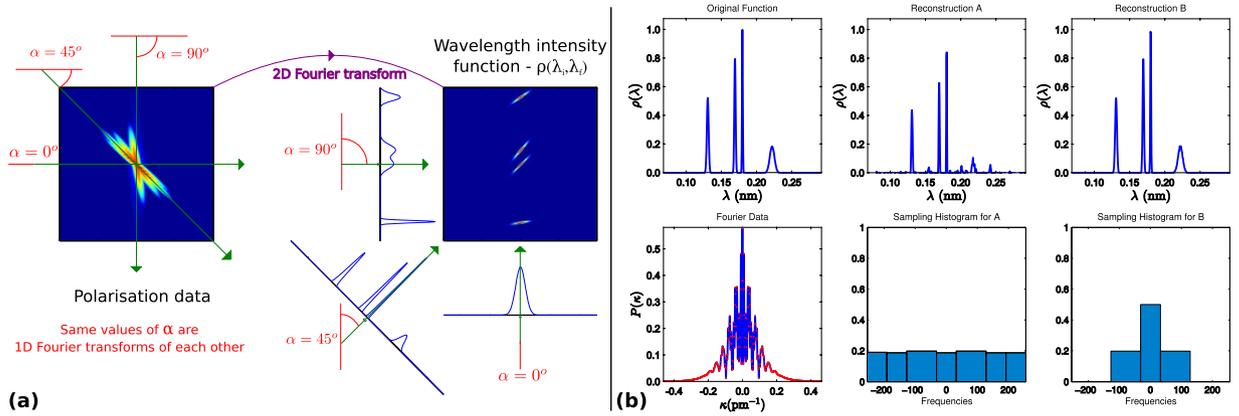
$$P(\boldsymbol{\kappa}) = \int \rho(\boldsymbol{\lambda}) e^{2\pi i \boldsymbol{\kappa} \cdot \boldsymbol{\lambda}} d\boldsymbol{\lambda}, \quad \boldsymbol{\lambda} = (\lambda_i, \lambda_f), \quad \boldsymbol{\kappa} = (\kappa_i, \kappa_f) \in \mathbb{R}^2. \quad (10)$$

Here  $\rho(\lambda_i, \lambda_f)$  denotes the *Wavelength Intensity Function* describing the distribution of He atoms that reach the detector according to initial and final wavelengths.

The change in wavelength can be caused by the creation or annihilation of surface phonons. A short description of the scattering upon surface vibrations can be found in the Supplementary Information. An example of a wavelength intensity function is displayed on the right-hand side of Fig. 2a showing four prominent features. Assuming the features in the plot originate from phonon phenomena on the crystal surface, the classification into creation/annihilation/elastic is according to the energy change  $\Delta E$ <sup>28</sup>:

$$\Delta E = E_f - E_i = \frac{\hbar^2}{2m}(k_f^2 - k_i^2) = \frac{\hbar^2}{2m} \left( \frac{1}{\lambda_f^2} - \frac{1}{\lambda_i^2} \right). \quad (11)$$

Looking at Fig. 2a as an example, we see that the key features of the wavelength intensity function can be broken down into lines of various slants which suggests that treating the function as one-dimensional would be advantageous. It is precisely because of this decomposition into slanted regions that the two-dimensional problem is often reduced to a one-dimensional one using the Fourier slice theorem: Instead of the two-dimensional  $(\lambda_i, \lambda_f)$



**Figure 2. Illustration of a typical wavelength intensity function and the Fourier slice theorem (left panel). The right panel shows a demonstration of the CS for a 1D example shown on the left. (a)** The wavelength intensity function (on the right) is the two-dimensional Fourier transform of the polarisation data (on the left). The Fourier slice phenomenon is demonstrated by the green lines on the right which indicate the direction of integration with the corresponding projections shown as lines in blue. The green lines on the left hand side represent the one-dimensional Fourier transforms of the projections shown on the right. The polarisation data intensity is shown on a log scale for the sake of readability. **(b)** Demonstrating CS for the 1D 45° projection shown in Fig. 2a, using uniform and multilevel sampling. Samples are taken from the Fourier data according to the sampling histograms shown. (Note that the frequency term in the sampling histogram relates to the values of  $\kappa$  at which  $P(\kappa)$  is sampled). Sampling pattern A is unreliable in reconstructing the rightmost feature as it is the least sparse of the four peaks while sampling pattern B remedies this by taking more of the lower frequency values that it depends upon. Reconstructions are at a resolution of 512 data points.

space one considers the projection onto a particular line defined by the angle  $\alpha$ . By restricting the measurement of  $P$  along the line  $\{(\kappa \cos \alpha, -\kappa \sin \alpha), \kappa \in \mathbb{R}\}$  the Fourier transform (10) becomes:

$$P(\kappa_i \cos \alpha, -\kappa_i \sin \alpha) = \int \rho_\alpha(\tau_1) \exp(2\pi i \kappa_i \tau_1) d\tau_1, \tag{12}$$

where  $\rho_\alpha(\tau_1)$  denotes the integral of  $\rho(\lambda)$  along the line  $\{R_\alpha(\tau_1, \tau_2): \tau_2 \in \mathbb{R}\}$ . For a more detailed derivation of the Fourier slice theorem please refer to the Supplementary Information.

Figure 2a shows how the Fourier slice theorem applies to the wavelength intensity function. Notice that different angles of integration produce different results, especially when it comes to discerning different features. Since we know beforehand that an elastic peak lies along the line  $\lambda_i = \lambda_j$  we expect that an integration angle of  $\alpha = \pi/4$  will produce the best results for resolving this feature as a single spike.

With this projection, we can treat the problem (12) as a one-dimensional version of (10) with a new wavelength intensity function  $\rho_\alpha(\lambda)$

$$P_\alpha(\kappa) = \int_a^b \rho_\alpha(\lambda) e^{2\pi i \kappa \cdot \lambda} d\lambda, \quad \lambda, \kappa \in \mathbb{R}. \tag{13}$$

**Compressed Sensing.** In this section we shall assume that we have already reduced the problem to one dimension and write  $P_\alpha, \rho_\alpha$  from (13) as  $P, \rho$ .

One can discretise (13) by breaking up the interval  $[a, b]$  into a uniform grid of  $N$  points  $\lambda_{i,N}$  as in (5), leading to the following matrix equation:

$$g_j = \sum_{l=-N}^N A_{j,l} f_l, \quad j = 0, \dots, 2N, \quad g_j = \tilde{\rho}_N(\lambda_{j,N}), \tag{14}$$

where  $f_l = \text{Constant}(l) \cdot P(l\varepsilon)$  and  $A$  is a DFT matrix.

Currently, to obtain the full vector  $(g_j)_{j=0, \dots, 2N}$  we need to know the entire vector  $(f_l)_{l=-N, \dots, N}$ . If we only had knowledge of a fraction of the entries of  $f$  we can no longer use (14) to determine  $g$  directly as the problem is now underdetermined. Therefore, the problem is not well posed and has to be modified.

The matrix equation (14) can be inverted to give

$$f_j = \sum_{l=0}^{2N} A_{j,l}^{-1} g_l, \quad j = -N, \dots, N. \tag{15}$$

Now suppose that  $\Omega \subset \{-N, \dots, N\}$  denotes the set of indices corresponding to the samples of  $f$  that are measured and  $P_\Omega f$  denotes the projection onto these samples. With this notation  $P_\Omega f$  denotes the vector of samples that are measured. Therefore when we subsample from  $\{-N, \dots, N\}$  equation (15) becomes

$$P_{\Omega}f = P_{\Omega}A^{-1}g. \quad (16)$$

The classical CS approach is to solve this problem via the now well established  $\ell^1$  recovery problem

$$\min_{g \in \mathbb{C}^N} \|Wg\|_{\ell^1} \quad \text{subject to} \quad P_{\Omega}f = P_{\Omega}A^{-1}g, \quad (17)$$

where  $W$  is some transformation that should make  $g$  sparse. Typically this is a wavelet transformation. We can then solve this kind of problem quickly and conveniently using convex solvers such as the (SPGL1) algorithm<sup>29</sup>.

The classical idea of CS is that  $\Omega$  should be chosen uniformly at random. In this case the number of samples  $m = |\Omega|$  must satisfy

$$m \gtrsim N \cdot \mu(A^{-1}W^{-1}) \cdot s \cdot \log(N), \quad (18)$$

in order to guaranty successful recovery with high probability, where  $\mu(B) = \max_{1 \leq i, j \leq N} |B_{ij}|^2$ . In the case where  $B = A^{-1}W^{-1}$  as above with any wavelet transform  $W$  we have that  $\mu(B) = 1$ . In this case, as well as many others, uniform random sampling may give suboptimal results and one has to sample with (structured) variable density sampling, see<sup>21</sup> and references therein. The key problem is that the optimality of variable density sampling depends on the signal itself<sup>21,24,25</sup>, and thus designing the best sampling pattern is a very delicate task. We will give a short demonstration below.

**How to do structured sampling.** The key to understanding structured sampling is to understand the structure of the signal. For example, the coefficients of a signal in a wavelet basis typically have a very specific level structure. This is known as sparsity in levels.

*Sparsity in levels.* Let  $x$  be a  $\mathbb{C}^N$  vector. For  $r \in \mathbb{N}$  let  $\mathbf{M} = (M_1, \dots, M_r) \in \mathbb{N}^r$  with  $1 \leq M_1 < \dots < M_r$  and  $\mathbf{s} = (s_1, \dots, s_r) \in \mathbb{N}^r$ , with  $s_l \leq M_l - M_{l-1}$ ,  $l = 1, \dots, r$ , where  $M_0 = 0$ . We say that  $x$  is  $(\mathbf{s}, \mathbf{M})$ -sparse if, for each  $l = 1, \dots, r$ ,  $\Delta_k := \text{supp}(x) \cap \{M_{l-1} + 1, \dots, M_l\}$ , satisfies  $|\Delta_k| \leq s_l$ . This known structure can be utilised when designing the sampling strategy and is the motivation behind multilevel sampling.

*Multilevel sampling.* Let  $r \in \mathbb{N}$ ,  $\mathbf{N} = (N_1, \dots, N_r) \in \mathbb{N}^r$  with  $1 \leq N_1 < \dots < N_r$ ,  $\mathbf{m} = (m_1, \dots, m_r) \in \mathbb{N}^r$ , with  $m_l \leq N_l - N_{l-1}$ ,  $l = 1, \dots, r$ , and suppose that  $\Omega_l \subseteq \{N_{l-1} + 1, \dots, N_l\}$ ,  $|\Omega_l| = m_l$ ,  $l = 1, \dots, r$ , are chosen uniformly at random, where  $N_0 = 0$ . We refer to the set  $\Omega = \Omega_{\mathbf{N}, \mathbf{m}} = \Omega_1 \cup \dots \cup \Omega_r$ , as an  $(\mathbf{N}, \mathbf{m})$ -multilevel sampling scheme. The key is that in the case of Fourier sampling, represented by the  $B$  above, combined with a wavelet transform  $W$  such that the recovery problem becomes (17), the multilevel sampling should match the level structure of the wavelets. More precisely,  $\mathbf{N} = \mathbf{M}$ . In this case, if  $x$  is  $(\mathbf{s}, \mathbf{M})$ -sparse with total sparsity  $s = s_1 + \dots + s_r$ , and  $s_1 = M_1 = m_1$ , then the total number of samples needed is

$$m = m_1 + \dots + m_r \gtrsim s_1 + (s_2 + s_3 + \dots + s_r) \log(N). \quad (19)$$

In particular, by utilizing the level structure in the sampling, one can outperform the standard CS results. For a more in-depth analysis and explanation see refs 21 and 25. See also refs 14 and 30 for early versions of this kind of sampling.

**Continuous Compressed Sensing.** The motivations behind continuous CS are: (i) to obtain a continuous approximation in the CS reconstruction, as opposed to a discrete approximation, as this allows for an easy non-linear change of variables to obtain the scattering function and the intermediate scattering function. (ii) If an alternative basis to the Fourier representation yields a better representation of the function to be recovered, one wants the freedom to use that. In particular, one can try approximating the wavelength intensity function  $\rho$  in terms of a new *Reconstruction Basis*  $\sigma_n$ ,  $n \in \mathbb{N}$ :

$$\tilde{\rho}(\lambda) \approx \sum_{n=1}^N \langle \rho, \sigma_n \rangle \sigma_n(\lambda), \quad \lambda \in [a, b]. \quad (20)$$

For technical reasons, one often requires these functions to form an orthonormal basis of  $L^2[a, b]$ , e.g. Legendre polynomials, splines, wavelets etc., although this condition can be relaxed to other groups of functions like frames<sup>31</sup>. For this paper we shall be using Daubechies wavelets<sup>32</sup> exclusively as our reconstruction basis. Let us quickly discuss *why* one would want to work with another basis.

Apart from the benefit of keeping the problem continuous, one also has the freedom to *choose* which basis  $\sigma_n$  to work with, making the approach more versatile than a straight DFT approach, where we are essentially forced to work with a pixel basis every time.

Furthermore, the notion of *sparsity* is now in terms of the coefficients  $\langle \rho, \sigma_n \rangle$ , which means we have the additional advantage of choosing a basis that makes the function  $\rho$  sparse. As we shall see, this opens up the possibility of using compressed sensing where traditional sparsity does not hold. In addition, this approach is closer to the philosophy that  $\rho$  being sparse should relate to  $\rho$  having low information content; a choice of basis  $\sigma_n$  that makes  $\rho$  sparse tells us how to (approximately) express the function  $\rho$  with a few non-zero coefficients.

Since one is still sampling data that corresponds to Fourier coefficients of  $\rho$ , it is impossible to exclusively work with their choice of basis  $\sigma_n$ . Instead one has to convert Fourier series coefficients into coefficients in the basis  $\sigma_n$ . This is achieved by working with the infinite change of basis matrix for the two bases:

$$B_{l,n} = \langle \sigma_n, \chi_{l,\varepsilon} \rangle, \quad n \in \mathbb{N}, \quad l \in \mathbb{Z}. \quad (21)$$

As opposed to (17), we now end up solving the infinite dimensional convex optimisation problem of finding

$$h^* \in \operatorname{argmin}_{h \in \ell^2(\mathbb{N})} \{ \|h\|_1 : P_\Omega B h = P_\Omega f \}, \quad (22)$$

Again  $P_\Omega$  denotes the projection onto the samples we have taken. In practice we cannot solve for the infinite solution to (22), therefore we truncate the reconstruction basis in a similar fashion to how we truncate the Fourier basis. This means we end up computing

$$h^* \in \operatorname{argmin}_{h \in \mathbb{C}^N} \{ \|h\|_1 : P_\Omega B P_N h = P_\Omega f \}, \quad (23)$$

where  $P_N$  denotes the projection onto the first  $N$  functions in the reconstruction basis. This problem is now numerically feasible since the submatrix  $P_\Omega B P_N$  is now finite (see<sup>21</sup> for estimates on how to choose  $N$ ). The solution to (23), let's say  $h^*$ , is recognised as the (approximate) wavelet coefficients of the intensity function. We can then use these wavelet coefficients to compute an approximation to  $\rho$  evaluated at any point on the interval  $[a, b]$  by following (20):

$$\rho(\lambda) \approx \sum_{n=1}^N h_n^* \sigma_n(\lambda), \quad \lambda \in [a, b]. \quad (24)$$

From here one can use the same multilevel sampling techniques as the discrete case to reconstruct  $(s, \mathbf{M})$ -sparse coefficients. Moreover, the sampling rule (19) also applies in this case. Note, however, that solving (23) is very different from solving (17). Indeed,  $B$  in (21) is an infinite matrix, moreover  $P_\Omega B P_N \neq P_\Omega A^{-1} W^{-1}$  where we recall  $A^{-1}$  and  $W^{-1}$  from (17). Both approaches, (23) and (17), are approximations to the true continuous problem (22), however, the discretisation is done differently. The discretisation in (17) turns out to be suboptimal, which can be seen as follows. As discussed in Compressing Spin-Echo Spectra, if  $g = Af$ , where  $f$  is a finite vector consisting of the samples of  $P$  from (1), then  $g$  is a vector that is a rasterised version of the truncated Fourier series approximation to  $\rho$  (the true solution we are seeking). Thus, if we apply a discrete wavelet transform  $W$  to  $g$ , we see that the result  $\tilde{h} = Wg$  is precisely the vector of the wavelet coefficients of the truncated Fourier series approximation to  $\rho$ . This means that whatever artefacts the truncated Fourier series approximation suffers from, such as Gibbs ringing, are also transferred to  $Wg$ . Hence, by solving (17) we will, at best, recover the wavelet coefficients of the truncated Fourier series that may be a suboptimal approximation. The wavelet coefficients of the truncated Fourier series are of little interest. What we want are the wavelet coefficients of the true solution  $\rho$ . This can be formulated via the infinite linear system  $\tilde{f} = Bh$ , where  $\tilde{f}$  is the infinite vector of all samples of the values  $P(l\varepsilon)$ ,  $l \in \mathbb{Z}$ , and  $h$  is the infinite vector of all the true wavelet coefficients of  $\rho$ . In comparison, the discretisation based on the DFT yields  $f = A^{-1} W^{-1} \tilde{h}$ , where  $\tilde{h}$  is the vector of wavelet coefficients of the truncated Fourier series. Note that (22) is an infinite-dimensional problem, thus, we need to approximate it via (23). However, as  $N$  grows, any solution to (23) converges to a solution to (22).

(23) is not based on the traditional DFT, however, it is possible to find a fast  $n \log(n)$  implementation for applying the matrix  $P_\Omega B P_N$ , the finite section of the infinite matrix  $B$ , to any vector. For further explanation and numerical examples demonstrating the benefits of the continuous approach and the differences with the discrete approach, see<sup>20–22</sup>.

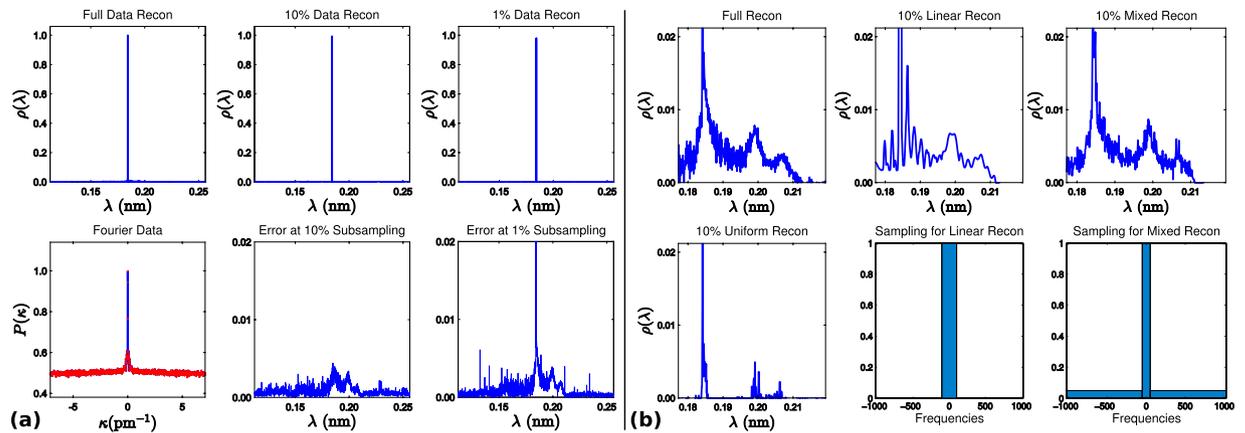
## CS for Phonon Detection

In this section we look at the performance of the CS approach described in the previous section by looking at examples of phonon detection. We shall first look at its effects on the one-dimensional projections shown previously and then focus on a real <sup>3</sup>He spectrum for scattering of gold where more exotic signal behaviour is present.

**Simulated 1D Example.** For consistency with previous sections we shall first work with the 45° projection shown earlier. Although this is a simplified model it clearly demonstrates some of the basic properties of compressed sensing. Reconstructions are shown in Fig. 2b.

Recall that we project along a 45° angle in an attempt to reduce the spread caused by the inaccuracy of the wavelength of the incident He<sup>3</sup> beam. In particular, since we know that there will always be an elastic feature in the wavelength intensity function which itself is slanted at 45°, this choice is seen as ideal for refocusing the various phonon features to be closer to that of a delta spike. Not only is this useful in preventing features from overlapping each other but this also increases sparsity which is ideal for compressed sensing; if features are sparse then by the rule (18) we can subsample to a great degree since the signal itself is very sparse. In this case, a change of basis may not be needed.

If one goes for a uniformly random approach to subsampling (as in Reconstruction A), as opposed to multilevel sampling discussed above, then there is only so far that one can go before problems occur. At around 20% subsampling, the reconstruction becomes unreliable in recovering the least sparse of the features on the far right. With 30% the rightmost feature is typically reconstructed but it is nonetheless unreliable. There is however an even more effective way of reliably reconstructing the rightmost feature by using multilevel sampling (as in Reconstruction B). The theory on how to design optimal structured multilevel sampling strategies is very new<sup>21,24,25</sup> and this is a highly unexplored topic. We do not attempt to seek optimality here, as this paper is about establishing the effectiveness of CS in HeSE. Note that, since there is no wavelet change of basis ( $W = I$ , the



**Figure 3. Compressed sensing reconstructions for a measured gold(111) phonon spectrum.**

(a) Compressed sensing reconstructions for a gold phonon spectrum (a.u.): With this choice of viewing range, only the elastic peak can be clearly seen. In this figure the sampling is performed uniformly at random which recovers the highly sparse central peak. This is suboptimal compared to the multilevel sampling in Fig. 3b. Frequencies sampled are from the range  $\{-1024, \dots, 1023\}$  and reconstructions are at a resolution of 2048. (b) Compressed sensing reconstructions for the same gold phonon spectrum as in Fig. 3a, zoomed in so that features beside the elastic peak are visible. Notice that the smaller features shown here are over 200 times smaller than the elastic peak. Reconstructions shown here are not only uniform (as in the bottom left graph) but also linear (i.e. straight Fourier series) and non-linear examples using roughly the same number of samples across each.

identity, in (17)) in this case, the theoretical understanding of the effect of multilevel sampling is not fully understood. This is still work in progress together with optimality conditions.

**Real Phonon Spectrum.** As we have already mentioned, real phonon spectra contain more exotic features than the simulation given in the previous example. Naturally noise adds to the data due to some experimental uncertainties of the measurement, but more unusual are the relative sizes and shapes of the various features.

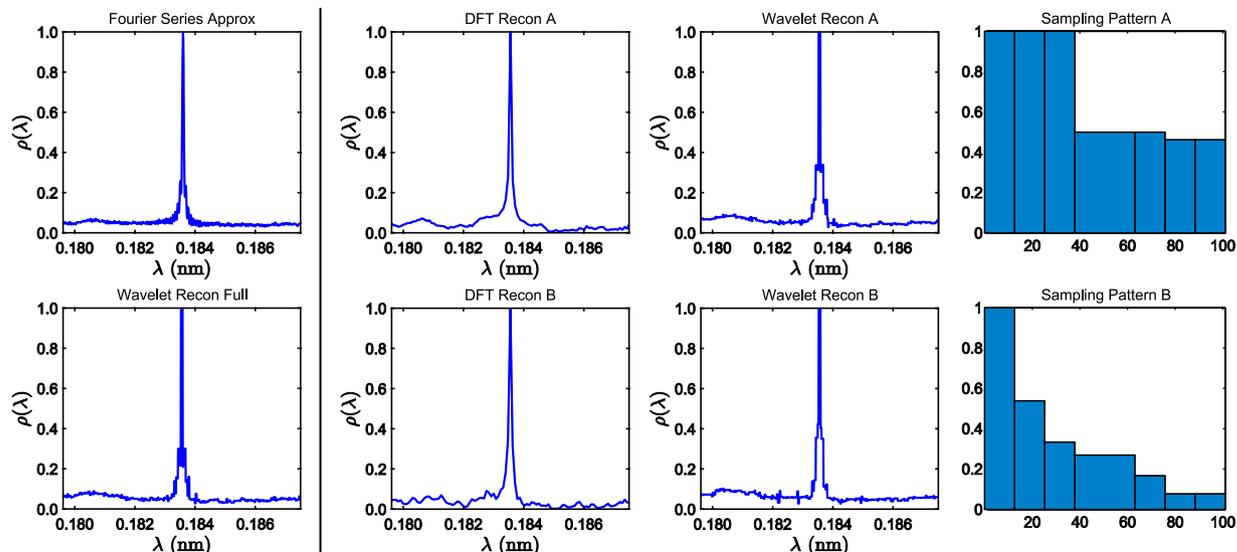
In Fig. 3a we have uniform sampling reconstructions for a typical gold(111) spectrum (for more details see Supplementary Information) with projection at  $45^\circ$  to focus on the elastic peak, which is the only clearly visible feature in the graphs. The peak is extremely fine and is in fact even smaller than the pixel resolution used for reconstruction (2048) which can be determined from the observation that the Fourier data has yet to decay to zero near the highest frequencies. Consequently this is an ideal situation for CS since this feature is almost as sparse as can be. Hence, one can subsample to a much greater degree (e.g. 1%) than in the previous example.

However, what has happened to the other inelastic features in this spectrum? At first one might come to the conclusion that they are not there at all, but, focusing on a small part of this spectrum reveals features that are over 200 times smaller than the measured intensity of the large elastic spike. Figure 3b shows various CS reconstructions zoomed in on this region. Notice that we still have the elastic peak visible, along with a couple of more (and less sparse) features. Like in the previous example, we expect the smooth features to be more dependent upon the lower frequency samples. Therefore, when we attempt to take just the first 10% of samples all from the lowest frequencies (the *linear* reconstruction) we find that these features are at the very least present, unlike the 10% uniform sampling approach where only the central peak remains. On the other hand, the central peak suffers from Gibbs artifacts which manifests themselves as wave like features near the central peak as well as broadening of the peak itself.

Instead, one can opt for a mix of these methods by taking the first 5% of samples from the lowest frequencies and the other 5% taken uniformly from the rest. This approach empirically performs the best out of the three methods, resolving the low resolution features without the Gibbs artefacts of the linear approach. Note that, as demonstrated in<sup>21,24</sup>, the optimal sampling procedure is signal structure dependent. How to choose optimal multilevel sampling is beyond the scope of this paper, and we have deliberately chosen a simple two-level sampling pattern, which is a reasonable all rounder, in order to demonstrate the effectiveness of the sampling technique.

**Comparing CS Techniques.** In this section we look at an example of how the continuous wavelet approach to compressed sensing can be used to tackle problems that are beyond the capabilities of the traditional compressed sensing approach described earlier.

Prior to data acquisition the spacing in current (equivalent to spacing in  $\kappa$ ) must be chosen, which in turn determines the length of the wavelength window  $[a, b]$  that the wavelength intensity function  $\rho$  is constructed over. If  $\rho$  is not truly supported on this window, then by (2) we instead reconstruct the periodised version of  $\rho$ . In particular, if peaks in the intensity function decay particularly slowly relevant to the window then the intensity function will stay considerably above zero throughout that window. Because of this, the traditional compressed sensing approach applied earlier cannot be used successfully here as the function is maximally non-sparse.



**Figure 4. (Continuous vs discrete CS) Reconstructions of a spin-echo spectrum (a.u.) of CoPc molecules deposited on Ag(001) with a noticeable baseline feature.** (See Supplementary Information for details on the sample preparation). In the DFT reconstructions the baseline level of around 0.1 is no longer flat leaving bumpy artefacts while the wavelet reconstructions preserve this flat feature. As we are only subsampling from 101 frequency points, considerable Gibbs artefacts are present in the Fourier series approximation. The DFT reconstructions have a resolution of 101 points, while the continuous Fourier series and wavelet reconstructions have been rasterised at a resolution ten times this number.

However, if one recalls the wavelet reconstruction bases that are used, then one quickly notices that they both have a constant function as the first basis function. This effectively means that the base level caused by slow decay is captured by this single basis function, which keeps the function sparse in these bases.

Figure 4 compares the two compression techniques for the diffusion of cobalt phthalocyanine (CoPc) on Ag(001) with an observable baseline feature. The full set of polarisation data points only corresponds to the first 101 frequencies and therefore there are noticeable Gibbs artefacts around the elastic peak in the Fourier series approximation. This strongly suggests that the Fourier series approximation here is not a particularly accurate approximation to the true intensity function.

Furthermore, the wavelet approximation aims to reconstruct the true underlying continuous wavelength intensity function, unlike the DFT approach which attempts to reconstruct a discretised form of the Fourier series approximation. Consequently, even with full sampling, the wavelet reconstruction is noticeably different to the Fourier series approximation. This reflects the fact that, as we are handling real data, we cannot directly compare to the true underlying wavelength intensity function.

Regardless we clearly observe that the baseline feature is preserved under subsampling using wavelets where the DFT approach clearly fails, matching predictions based on sparsity observations earlier. Note that both techniques use exactly the same samples. While we are able to subsample to a reasonable degree here ( $\approx 33\%$ ), one should ideally work with a larger range of frequencies to truly exploit the benefits of this approach, i.e. subsampling from polarisation data with thousands of points rather than hundreds.

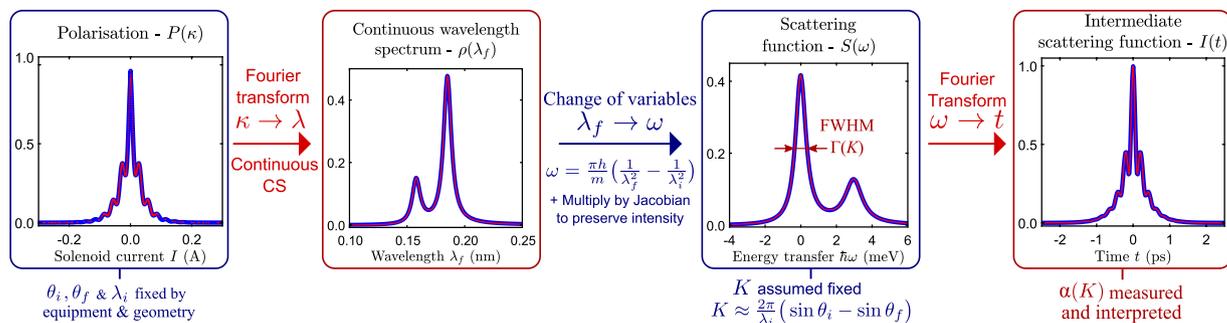
**Surface Diffusion Using Continuous CS.** Surface diffusion, in contrast to phonon studies, deals with species (atoms, molecules) adsorbed on surfaces which diffuse over large distances. It aims to identify the main mechanisms governing the motion on surfaces (i.e. the energetic landscape governing the adsorbate dynamics) and, therefore, to characterize their diffusive behaviour.

The Van Hove formalism of correlation functions in time and space<sup>11</sup>, initially developed for neutron scattering (see Sec. S.5 in the Supplementary Information for more details), provides us with a theoretical formalism to understand and interpret the experimentally determined ISF  $I(\mathbf{K}, t)$  in terms of diffusive regimes. (Here we have omitted the  $\Delta$  in the notation of the momentum transfer  $\Delta\mathbf{K}$  for the sake of simplicity.) The Van Hove correlation contains all the information related to the dynamics of particles in the system but it cannot be directly measured. However, its properties can be inferred from partial knowledge of the SF/ISF (see Fig. 5). In particular, different diffusive regimes can be identified through the *dephasing rate*  $\alpha(\mathbf{K})$ , which describes the decay in time  $t$  of the ISF as a function of  $\mathbf{K}$ <sup>33</sup>:

$$\alpha(\mathbf{K}) = \inf\{t > 0: |I(\mathbf{K}, t)| = e^{-1}|I(\mathbf{K}, 0)|\}. \quad (25)$$

Its Fourier pair is the so-called *quasi-elastic broadening*  $\Gamma(\mathbf{K})$  and is defined as the half-width-half-maximum of the quasi-elastic profile in the SF, i.e. in the peak centred around  $\Delta E = 0$  of the SF in Fig. 5.

Therefore, it is essential to determine the “true” ISF from the measured polarisation  $P(\kappa)$ . The processing of the data from  $P(\kappa)$  to the ISF requires to go through a series of Fourier transformations and change of variables as



**Figure 5.** A detailed 1D description of the full loop presented in Fig. 1. Blue lines denote the true underlying signal and red lines denote the reconstructions (except for the left panel where the red dots denote the sampling points). Sampling points are taken according to the sampling histogram present in Fig. 1. Note that there is no comparison with standard discrete CS here as the change of variable technique is impossible in the discrete setup. This can only be done in the continuous case.

shown in Fig. 5. The drawback is that the change from wavelength to energy in (11) is incompatible with the standard DFT based CS methods because the uniform grid of coordinates in wavelength space is transformed onto a non-uniform grid in  $(\mathbf{K}, \Delta E)$ -space (see Sec. S.6 in the Supplementary Information for more details on this variable transform). With continuous CS one can work backwards, first specifying a uniform grid in  $(\mathbf{K}, \Delta E)$ -space which is converted to a non-uniform grid in wavelength space. As a result, since the reconstructed solution of (22) is a function rather than a vector, one can directly sample the wavelength intensity function on this non-uniform grid and compute the SF on a uniform energy grid. This allows to apply the standard DFT algorithm to obtain the final ISF.

**Measuring Diffusion and Phonons in One Go.** Through the different sections of this paper we have seen that the HeSE technique can explore dynamical processes spanning from phonon measurements in the sub-picosecond time window up to diffusion in the cents of picoseconds scale. Yet the continuous CS approach enables us not only to reduce the measurement times, it is also a powerful method to circumvent the experimental complications when trying to capture very different dynamical processes within a single set of data.

Firstly, the capability to reconstruct the spectrum out of a subsampled set of measured currents allows to cover a very large current window with a high resolution in a reasonable measurement time. This provides us with a very broad and highly resolved energy window where we can separate accurately the phonon (the inelastic peak in the spectrum) from the diffusive contribution (the quasi-elastic peak centered at  $\Delta E = 0$  meV). Secondly, the flexibility of the continuous CS approach allows us to evaluate the reconstructed signal in any desired wavelength grid which greatly simplifies the application of the standard DFT algorithm linking the last two stages of the data processing, from the SF to the ISF in Fig. 5.

Hence the CS approach provides us with a flexible framework where the different physical processes contained in a single data set can be disentangled and analysed in the energy domain (where inelastic features such as phonons can be easily investigated) or in the time domain (where diffusive processes are easier to characterise).

## Conclusion and Summary

We have used compressed sensing (CS) techniques to measure and reconstruct surface dynamical processes with a helium spin-echo spectrometer. This work demonstrates that in particular the continuous CS approach can be used to reduce measurement times by at least an order of magnitude whilst capturing both phonon and diffusion processes simultaneously.

Continuous CS allows us to reconstruct spectra with a very broad and highly resolved energy window in a reasonable measurement time so that we can separate accurately phonon events and diffusion processes measured in a single data set. Since the continuous CS approach allows to evaluate the reconstructed signal on any desired grid this makes it easy to switch between the energy and the time domain and to analyse the underlying physical processes. Not only has this made current helium spin-echo experiments more convenient, but this has also brought forward future projects that were originally deemed too time-consuming to measure. Eventually the final goal is to capture the entire scattering function over all of  $(\mathbf{K}, \Delta E)$ -space using a two-dimensional continuous CS method.

The developed methodology can also be applied to other scattering frameworks and the authors hope that these advances will be quickly brought to the attention of the neutron scattering and X-ray communities.

## References

1. Farias, D. & Rieder, K.-H. Atomic beam diffraction from solid surfaces. *Rep. Prog. Phys.* **61**, 1575 (1998).
2. Benedek, G. & Toennies, J. P. Helium atom scattering spectroscopy of surface phonons: genesis and achievements. *Surf. Sci.* **299**, 587–611 (1994).
3. Tamtögl, A. *et al.* Graphene on Ni(111): Electronic Corrugation and Dynamics from Helium Atom Scattering. *J. Phys. Chem. C* **119**, 25983–25990 (2015).
4. Jardine, A. P., Hedgeland, H., Alexandrowicz, G., Allison, W. & Ellis, J. Helium-3 spin-echo: Principles and application to dynamics at surfaces. *Prog. Surf. Sci.* **84**, 323–379 (2009).
5. Alexandrowicz, G. & Jardine, A. P. Helium spin-echo spectroscopy: studying surface dynamics with ultra-high-energy resolution. *J. Phys.: Condens. Matter* **19**, 305001 (2007).

6. Mezei, F. Neutron Spin Echo: Proceedings of a Laue-Langevin Institut Workshop Grenoble, October 15–16, 1979, chap. The principles of neutron spin echo, 1–26 (Springer, 1980).
7. Jardine, A. P., Alexandrowicz, G., Hedgeland, H., Allison, W. & Ellis, J. Studying the microscopic nature of diffusion with helium-3 spin-echo. *Phys. Chem. Chem. Phys.* **11**, 3355–3374 (2009).
8. Hedgeland, H. *et al.* Measurement of single-molecule frictional dissipation in a prototypical nanoscale system. *Nat. Phys.* **5**, 561–564 (2009).
9. Harten, U., Toennies, J. P. & Woll, C. Helium Time-of-flight Spectroscopy of Surface-phonon Dispersion Curves of the Noble Metals. *Faraday Discuss. Gem. Soc.* **80**, 137–149 (1985).
10. Kole, P. R., Jardine, A. P., Hedgeland, H. & Alexandrowicz, G. Measuring surface phonons with a <sup>3</sup>He spin echo spectrometer: a two-dimensional approach. *J. Phys. Condens. Matter* **22**, 304018 (2010).
11. Hove, L. V. Correlations in space and time and born approximation scattering in systems of interacting particles. *Phys. Rev.* **95**, 249 (1954).
12. Candès, E. J., Romberg, J. & Tao, T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **52**, 489–509 (2006).
13. Candès, E. J. An introduction to compressive sensing. *IEEE Signal Process. Mag.* **25**, 21–30 (2008).
14. Donoho, D. L. Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
15. Lustig, M., Donoho, D. L., Santos, J. M. & Pauly, J. M. Compressed sensing mri. *IEEE Signal Process. Mag.* **25**, 72–82 (2008).
16. Guerquin-Kern, M., Häberlin, M., Pruessmann, K. & Unser, M. A fast wavelet-based reconstruction method for magnetic resonance imaging. *IEEE Trans. Med. Imag.* **30**, 1649–1660 (2011).
17. Holland, D. J., Bostock, M. J., Gladden, L. F. & Nietlisbach, D. Fast multidimensional nmr spectroscopy using compressed sensing. *Angew. Chem. Int. Ed.* **50**, 6548–6551 (2011).
18. Galvis-Carreño, D. F., Meijia-Melgarejo, Y. & Arguello-Fuentes, H. Efficient reconstruction of raman spectroscopy imaging based on compressive sensing. *DYNA* **81**, 116–124 (2014).
19. Andrade, X., Sanders, J. N. & Aspuru-Guzik, A. Application of Compressed Sensing to the Simulation of Atomic Systems. *Proc. Natl. Acad. Sci. USA* **108**, 13928–13933 (2012).
20. Adcock, B. & Hansen, A. Generalized sampling and infinite dimensional compressed sensing. *Found. Comp. Math.* 1–61 (2015).
21. Adcock, B., Hansen, A. C., Poon, C. & Roman, B. Breaking the coherence barrier: A new theory for compressed sensing arXiv:1302.0561 (2014).
22. Adcock, B., Hansen, A. C., Roman, B. & Teschke, G. Generalized sampling: stable reconstructions, inverse problems and compressed sensing over the continuum. *Adv. in Imag. and Electr. Phys.* **182**, 187–279 (2014).
23. Scheuer, J. *et al.* Accelerated 2D magnetic resonance spectroscopy of single spins using matrix completion. *Sci. Rep.* **5**, 17728 (2015).
24. Roman, B., Adcock, B. & Hansen, A. On asymptotic structure in compressed sensing, arXiv:1406.4178 (2014).
25. Adcock, B., Hansen, A. C. & Roman, B. *Compressed Sensing and its Applications: MATHEON Workshop 2013*, chap. The Quest for Optimal Sampling: Computationally Efficient, Structure-Exploiting Measurements for Compressed Sensing, 143–167 (Springer, 2015).
26. Vetterli, M., Marziliano, P. & Blu, T. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.* **50**, 1417–1428 (2002).
27. Markovich, T. *et al.* More accurate and efficient bath spectral densities from super-resolution arXiv:1307.4407 (2013).
28. McIntosh, E. M. *et al.* Measurement of the phason dispersion of misfit dislocations on the au(111) surface. *Phys. Rev. Lett.* **110**, 086103 (2013).
29. Berg, E. & Friedlander, M. Probing the pareto frontier for basis pursuit solutions. *SIAM J. Sci. Comput.* **31**, 890–912 (2008).
30. Studer, V. *et al.* Compressive fluorescence microscopy for biological and hyperspectral imaging. *Proc. Natl. Acad. Sci. USA* **109**, E1679–E1687 (2011).
31. Mallat, S. *A wavelet tour of signal processing* (Elsevier Academic Press, 2008), 3 edn.
32. Daubechies, I. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* **41**, 909–996 (1988).
33. Bée, M. *Quasielastic Neutron Scattering*, chap. Long-range Translational Diffusion, 148–175 (CRC Press, 1988).

## Acknowledgements

A. Jones acknowledges EPSRC grant EP/H023348/1, A. Tamtögl acknowledges support by the FWF (project J3479-N20), I. Calvo-Almazán acknowledges support from the Ramón Areces Institution and A. Hansen acknowledges support from a Royal Society University Research Fellowship as well as EPSRC grant EP/L003457/1. The authors would like to thank W. Allison for many helpful discussions.

## Author Contributions

A.J. implemented the main parts of the compressed sensing and wavelet reconstruction algorithms and A.T. performed the experimental measurements. I.C.-A. implemented the post-processing of the measured data, in particular the transformation steps from the measured data up to the intermediate scattering function. A.J., I.C.-A. and A.T. analysed the data and prepared the figures. All authors were involved in preparing and revising the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Jones, A. *et al.* Continuous Compressed Sensing for Surface Dynamical Processes with Helium Atom Scattering. *Sci. Rep.* **6**, 27776; doi: 10.1038/srep27776 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>